DOCUMENT RESUME

ED 474 881 UD 035 592

AUTHOR	Kane, Thomas J.; Staiger, Douglas O.; Geppert, Jeffrey
TITLE	Assessing the Definition of "Adequate Yearly Progress" in the House and Senate Education Bills.
PUB DATE	2001-07-15
NOTE	24p.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Accountability; Diversity (Student); *Educational Legislation; Elementary Secondary Education; Federal Legislation; Public Education; *Scores; Student Evaluation
IDENTIFIERS	North Carolina; Texas

ABSTRACT

In 2001, the U.S. House of Representatives and U.S. Senate both passed education bills with tough school accountability provisions. Both bills require states to test all students in grades 3-8 within 3 years and to separately report performance of subgroups (including racial and ethnic subgroups) within each school. An important innovation in both bills is the definition of "adequate yearly progress." This paper evaluates the implications of these two pieces of legislation for schools in North Carolina and Texas, two states with rapid increases in test scores between 1994-99. Results indicate that both bills ignore the natural volatility in school test scores by requiring increases in a school's test performance each year. Virtually every school in North Carolina and Texas would have failed to achieve "adequate yearly progress" at least once between 1994-99 under either the House or the Senate bill. By making the achievement of "adequate yearly progress" contingent on the improvements of each and every subgroup of students in a school, both measures disadvantage schools containing more than one racial or ethnic group. Recommendations include: pool performance over multiple years, maintain state flexibility to define adequate yearly progress, and do not penalize racially diverse schools. (SM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 474 881

Assessing the Definition of "Adequate Yearly Progress" in the House and Senate Education Bills

by
Thomas J. Kane, Hoover Institution and UCLA
Douglas O. Staiger, Dartmouth College
and
Jeffrey Geppert, National Bureau of Economic Research

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
T. Kane
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

July 15, 2001

Contact Information:
Thomas J. Kane
Professor of Policy Studies and Economics
School of Public Policy and Social Research
UCLA
3250 Public Policy Building
Los Angeles, CA 90095-1656
(310) 825-9413
tomkane@ucla.edu

UD 035 592

BEST COPY AVAILABLE

1. Introduction

This summer, the U.S. House of Representatives and the U.S. Senate have each passed education bills with tough school accountability provisions. Both bills require states to test all students in grades 3 through 8 within three years and to separately report the performance of subgroups of students, including racial and ethnic subgroups, within each school. However, the most important innovation in both bills is the definition of "adequate yearly progress." Congress has chosen to specify detailed test score expectations for schools and for subgroups of students within schools (including groups identified by race and ethnicity). Moreover, both bills would require states and school districts to intervene in schools that fail to meet those standards– initially by offering their students public school choice options and, eventually, by imposing more serious sanctions, such as reorganizing the school as a public charter school. The Administration hopes to have the measure signed into law by the end of the summer. Yet, neither bill's definition of "adequate yearly progress" has been subjected to careful scrutiny. In this paper, we evaluate the implications of both pieces of legislation for schools in North Carolina and Texas-- two states with rapid increases in test scores between 1994 and 1999.

The new federal requirements would arrive at a time when many states are experimenting with school accountability systems. By the spring of 2000, forty states had begun using student test scores to rate school performance. Twenty states are going a step further and attaching explicit monetary rewards or sanctions to a school's test performance. For example, California plans to spend $667 million on teacher and school incentives this year, providing bonuses of up to $25,000 to teachers in schools with the largest test score improvements. Some states (such as California) reward annual changes in a school's mean test score; while other states (such as North Carolina) reward states based upon value-added measures. By focusing on annual increase in the "percent proficient" in reading and in math, the proposed federal law is not always consistent with the accountability systems states are implementing.

We report several important findings:

Both bills ignore the natural volatility in school test scores, by requiring increases in a school's test performance each year. A school's mean test score will naturally fluctuate, depending upon the particular group of children being tested in a given year. Given that the average elementary school contains 68 children per grade level, test scores for a given grade or school can be sensitive to the talents or rowdiness of a particular cohort of children. For instance, in North Carolina, the proportion of 3rd through 5th grade students scoring at the "proficient" level grew by 3 percentage points per year in math and 2 percentage points per year in reading. However, the string of increases at the state level was not reflected in every school in every year: less than 2 percent of the elementary schools in the state witnessed a positive increase in proficiency in reading and in math for 5 straight years between 1994 and 1999.

Virtually every school in North Carolina and Texas would have failed to achieve "adequate yearly progress" at least once between 1994 and 1999 under either the House or Senate bills. Since schools failing to make "adequate yearly progress" are expected to provide school improvement plans, the law would have generated a large amount of paperwork for schools and the district administrators charged with evaluating those plans. However, since virtually every school would be required to do so, it is difficult to imagine that districts or states would have the resources to review those plans carefully. Moreover, under the House bill, roughly 96 percent of the schools in North Carolina and in Texas would have faced corrective action and three-quarters or more would have faced restructuring during those 5 years-- even though both states were experiencing rapid test score growth over that period.

By making the achievement of "adequate yearly progress" contingent on the improvements of each and every subgroup of students in a school, both measures disadvantage schools containing more than one racial or ethnic group. For example, among elementary schools in North Carolina, we estimate that eliminating the subgroup rules would almost double the annual passage rates for schools with 2 racial subgroups and would triple annual passage rate for schools with 3 racial subgroups. The problem is a statistical one, due to the independent fluctuations in scores for each group. When one group's scores are up, another group's scores are often down. There is no evidence that minority youth were falling behind white non-Hispanic youth in the diverse schools. Indeed, the growth in proficiency was slightly higher for black and Hispanic youth than for white youth in diverse schools. Morever, there is no evidence in Texas that the scores of Latino or African American youth grew any more rapidly in schools where their subgroup scores were counted separately than in schools where their scores did not count separately.

We close with some thoughts on how the legislation could be improved.

2. Evidence from North Carolina and Texas

Even when a school is on the right track, the path to improved student performance is rarely a straight path. Each two steps forward is often followed by one step back. The cause is often not a lack of resolve among school administrators or a waning desire among teachers and students. Rather, it is the natural fluctuation in performance that comes with the passing of successive cohorts of children through a school. Even if school performance is on an upward trend, the underlying rate of improvement can be temporarily dwarfed by the effect of having 5 really bright kids in a class one year and only 3 the next, or having a particularly rowdy group of friends together one year and not the next. Such volatility is a particular problem in elementary schools, since there are only 68 kids per grade level

in the average school nationally.[1] When there are so few kids in a class, a few stars or a few class clowns can generate large fluctuations in mean test scores. When looking for signs of improvement at the school level, one typically has to look to trends over several years, rather than the change in any single year.

In North Carolina between 1994 and 1999, the proportion of students in grades 3 through 5 scoring at the "proficient" level or higher in mathematics rose from 55 percent to 70 percent– roughly 3 percentage points per year. The proportion of 3rd through 5th grade students scoring at the proficient level in reading grew from 61 to 70 percent, or nearly 2 percentage points per year. Progress of that magnitude has made North Carolina the envy of many other states.

More than two-thirds of schools experienced an increase in math proficiency in the average year (68 percent) and just under two-thirds experienced an increase in reading proficiency in the average year (63 percent). However, both the House and Senate bills require increase in both proportions in a given year.[2] As one might expect, given that nearly all students take both tests, any improvements or declines in math or reading proficiency are related, but not perfectly. Only slightly more than half (51 percent) of the schools witnessed an increase in both math or reading proficiency in any given year.[3]

Table 1 reports the number of years between 1994 and 1999 that North Carolina elementary schools experienced positive increases in the proportion of students proficient in reading, math and in both subjects. Only 11 percent of schools witnessed an increase in math proficiency for 5 straight years, and only 6 percent witnessed an increase in reading proficiency for 5 straight years. However, less than two percent of schools witnessed an increase in both subjects for 5 straight years. Rather, it was most common for schools to have seen 3 years of increases and 2 years of declines over these 5 years. Indeed, 36 percent of schools experienced such a pattern.

---

[1]In North Carolina, the average elementary school had test scores for 218 students across grades 3 through 5, or 72 students per grade level. In Texas, the average elementary school was slightly larger, with 233 students with test scores in grades 3 through 5, or 78 students per grade level.

[2]Imagine 3 different proportions: The proportion of youth proficient in reading, P(R), the proportion of youth proficient in math, P(M) and the proportion of youth proficient in reading and math P(R and M). Both the House and Senate bills require increases in P(R) and in P(M) separately. We have experimented with rules built around P(R and M), which reduces some of the volatility.

[3]If changes in math proficiency were uncorrelated with changes in reading proficiency, we would have expected even fewer schools (43 percent) to have witnessed improvements on both tests in a given year (.683*.634=.43).

Achieving Adequate Yearly Progress under the House and Senate Bills

We use test results between 1994 and 1999 in North Carolina and Texas to estimate the likely implications of the House and Senate definitions of adequate yearly progress on school performance. An important feature of both bills is the use of subgroup targets. Our data allowed us to identify up to 6 subgroups within each school in North Carolina: African American (non-Hispanic) students, Asians, Native Americans, Hispanics, white non-Hispanics and students with limited English proficiency. All but one of the groups are mutually exclusive, with students from the Limited English Proficiency group spread among the other groups. Although the legislation does not specify a minimum sample size for each group to receive separate consideration, we required a subgroup to contain 15 students in order to achieve subgroup status. Because we could not identify students receiving free or reduced price lunches or disabled students every year, we did not allow for separate subgroup targets for these groups, even though both the House and Senate bills would have. As a result, our estimates should be understood as conservative, and probably overstate the proportion of schools making adequate yearly progress.

North Carolina and Texas both already test students each year from grades 3 through 8. As a result, in our estimates, we pool data from three grade levels (grades 3 through grade 5) when following the progress of elementary schools. However, many states currently test only one grade level in elementary schools. For instance, Massachusetts administers their state exam to students in 4th, 8th and 10th grades-- meaning that they currently test one grade level in elementary schools, one grade level in middle schools and one grade level in high schools. Although states would be required to test all grades (from grades 3 through 8) within three years, the calculation of adequate yearly progress would begin immediately. To the extent that adding additional grade levels would dampen the annual fluctuations, this is a second reason why our estimates probably overstate the proportion of schools achieving adequate yearly progress, since we begin with three grade levels per elementary school.

The Senate bill requires a one percentage point increase for every subgroup in the percentage of students proficient in math as well as in reading. In contrast, the House bill requires an annual increase in each subject sufficient to keep a school on track to achieve 100 percent proficiency at the end of twelve years. In other words, if 40 percent of African American students were proficient in math and 52 percent were proficient in reading in 1994, a school would need to achieve a 5 percentage point increase in math proficiency and a 4 percent increase in reading proficiency in order to achieve adequate yearly progress as defined by the House bill ((100-40)/12=5 and (100-52)/12=4). Because all groups averaged less than 88 percent proficiency in 1994, the House bill presents a higher hurdle than the Senate bill for the vast majority of schools.

The Senate bill also allows states to calculate the proficiency in a given year by averaging proficiency over 3 years. We calculated adequate yearly progress with and without allowing for 3-year rolling averages under the Senate plan. Because our data started in 1994, we were able to calculate 3-year rolling averages for 1996 through 1999. However, for 1994, we used the single year of data and,

for 1995, we used the average of 1994 and 1995. Unless states already had test score data for two prior years, our estimates would likely reflect the process that would be used upon initial implementation. (In the Appendix, we compare results generated by the House and Senate bills for the 3 years for which we could use 3 year rolling averages for every year.)

Table 2 reports the number of years in which NC elementary schools achieved adequately yearly progress between 1994 and 1999. All of the NC elementary schools would have failed the House definition of adequate yearly progress in at least one year between 1994 and 1999. Without averaging, all of the elementary schools also would have failed at least once under the Senate plan. Using 3-year rolling averages, 98 percent of schools would have failed at least once under the Senate plan.

However, schools were much more likely to have repeated failures using the House definition: 97 percent would have failed in 2 consecutive years over 5 years and 83 percent would have failed in 3 consecutive years. Using 3-year rolling averages under the Senate plan, 88 percent would have failed in 2 consecutive years and 62 percent would have failed in 3 consecutive years.

Sanctions

Both bills require those schools failing to make adequate yearly progress in any year to submit school improvement plans. Since virtually all elementary schools would have failed to make adequate yearly progress at least once within 5 years, both bills would have implied a large amount of paperwork at the school, district and state levels to produce, evaluate and respond to school improvement plans.

After one year of failure, the House bill also requires that students be given the option of attending another public school in the district. However, students are only allowed to transfer to schools which had achieved adequate yearly progress. Ironically, since 86 percent of schools would have failed to achieve adequate yearly progress in any given year, those students who did qualify for public school choice would have had few other options from which to choose. If the student were able to find a school to attend, the House plan would require the district to pay for transportation expenses. (Public school choice begins after 2 years of failure under the Senate plan, but a district would only have to pay for transportation expenses only after 3 years of failure.)

The more serious implications begin only after a school has failed to make adequate yearly progress in two or more consecutive years. Under either bill, there are two levels of sanctions after a school has failed to achieve adequate yearly progress in a given year: "corrective action" and "restructuring" (or, as it is referred to in the Senate bill, "reconstitution"). When a school falls under "corrective action" status, a district is required to offer Title I eligible students the option to use a portion of the school's federal funding to pay for tutoring or other supplemental educational services. In addition, the district is required to take one of several actions, such as replacing relevant school staff, implementing a new curriculum (along with the requisite teacher training) or re-opening the school as a

charter school. For schools that reach the last stage, restructuring status, the district is required to take even more serious steps, to either reopen the school as a charter school, turn the school over to a private entity, replace a majority of the school staff or allow the school to be taken over by the state.

"Corrective action" status is reached after 2 consecutive years of failure under the House plan and 3 consecutive years of failure under the Senate plan.[4] Our estimates suggest that 97 percent of North Carolina schools would have faced corrective action between 1994 and 1999 under the House plan and nearly two-thirds would have faced corrective action under the Senate plan (62 percent without averaging and 61 percent with averaging).

"Restructuring" is reached after 3 years of failure under the House plan and after 5 years of failure under the Senate plan. Our estimates suggest that 83 percent of North Carolina schools would have faced restructuring under the House plan within the first 5 years and a quarter of schools would have faced restructuring under the Senate plan. Presumably, these percentages would increase over time, since a school would have had to fail every year during the 5-year period we observed in order to qualify for "restructuring" under the Senate bill.

Adequate Yearly Progress in Texas, 1994-99

Using data available from the Texas Education Agency, we repeated the above exercise for Texas elementary schools between 1994 and 1999. In the Texas data, we were able to identify up to 4 subgroups within each school: white non-Hispanic youth, black non-Hispanic youth, Hispanic youth and economically disadvantaged youth. Our data reported the proportion of all students in a school and the proportion of each subgroup in the school that were proficient in reading and in mathematics. As we did with the North Carolina data, we assumed that a group had to contain 15 or more students in order to be counted separately as a subgroup. Over this period, Texas schools, like North Carolina schools, were achieving large increases in proficiency. However, as reported in Table 3, Texas schools would have fared little better than North Carolina schools if the proposed federal legislation had been in effect between 1994 and 1999. Despite making rapid gains, nearly every elementary school in Texas would have failed to make adequate yearly progress at least once over 5 years, under either the House or the Senate rules. Moreover, under the House rules, 96 percent of schools would have faced corrective action and 73 percent would have faced restructuring over those 5 years. Under the Senate rules, more than half of schools would have faced corrective action (with or without the 3-year rolling averages) and nearly a quarter would have faced restructuring, after failing every year for 5 years.

---

[4]This is a conservative interpretation of the language in the Senate bill. Given that it takes two years of achieving adequate yearly progress to emerge from "needs improvement" status, one reading of the Senate bill would have schools falling into corrective action status by simply by failing every other year.

Isolating the Impact of the Subgroup Rules

Both the House and Senate plans require schools to achieve improvements in test scores for all racial subgroups in order to achieve adequate yearly progress. The rules are intended to encourage schools to find ways to improve performance of all students and not to ignore disadvantaged minority students. However, because each subgroup's scores are bouncing around from year-to-year depending upon the particular collection of students being tested, such rules put diverse schools-- those with more than one racial or ethnic subgroup-- at a distinct disadvantage. When one group's test scores are up, another group's scores are often down and vice versa. Schools with multiple subgroups are much more likely to fail to make adequate yearly progress than schools with only one. Ironically, to the extent that disadvantaged minority students are more likely to attend schools with multiple subgroups, such rules may end up harming their intended beneficiaries.

Table 4 reports the proportion of schools achieving adequate yearly progress using the Senate rules for the years 1997 through 1999, when we would have had a sufficient number of years to calculate 3-year rolling averages for all groups. The data are reported by the average number of subgroups each school had during those years. As reported in the top panel of Table 4, schools with 2 subgroups were more than twice as likely to fail all three years than those with 1 subgroup. Those with 3 subgroups, were more than 4 times as likely to fail all three years.

However, the overall growth in test scores was not dramatically different in the racially homogeneous and racially diverse schools. The second panel in Table 4 reports the proportion of schools achieving adequate yearly progress if there were no subgroup rules (that is, if the only requirement were that each school achieve a 1 percentage point increase in the proportion of all students achieving proficiency.) As expected, the change in rules would have little effect on the schools with only 1 subgroup.[5] However, the change in rules would have resulted in a dramatic increase in the passing rate for diverse schools. For example, the proportion of schools achieving adequate yearly progress in an average year triples for schools with 3 subgroups from .15 to .45 and rises from .34 to .54 for schools with 2 sugroups.

Disadvantaged minority students are not being left behind in the more diverse schools. The bottom panel of Table 4 reports the difference in average annual growth in math proficiency between white and African American students and between white and Latino students, when a given school contained both groups. There were very small differences in the average growth rates by race. Three out of four differences is less than 1 percentage point. Moreover, in every case, the differences imply

---

[5]The results in the second panel of Table 4 differ from the results in the first panel for two reasons: first, a few schools averaging 1 subgroup of students over the period 1994-99 had more than 1 subgroup in a given year; second, even the schools with one subgroup of students may have less that 15 students in some other racial subgroups.

faster growth rates for minority students than for white students.

Then what accounts for the fact that diverse schools fare so poorly under the Senate and House plans? Each group's scores vary from year to year depending upon the specific group of individuals being tested. Because each group's scores vary in a random way, they are very weakly correlated with each other. Figure 1 portrays the change in the percentage of white non-Hispanic students who were proficient between 1998 and 1999 by the change in the percentage of African American students who were proficient (using the 3-year rolling average proficiency rate for each group in each year). Two facts are apparent: First, many schools had very large increases or large decreases in scores for either racial group between 1998 and 1999. Second, there was only a very weak relationship between the change in percent proficient for blacks and the change in the percent proficient for whites. The schools with large increases for one group did seem to be slightly more likely to have large increase for both groups, but only slightly.

Because of the volatility in test scores from year to year, requiring racially diverse schools to achieve targets for every subgroup is analogous to having them flip a coin twice each year and get heads every time. The table at the bottom of Figure 1 portrays the proportion of schools achieving more than a 1 percentage point growth in proficiency for different racial groups in different years. (The table was limited to the 677 elementary schools in North Carolina that had more than 15 black and more than 15 white students in 1998 and 1999.) More than two thirds (69 percent) of the schools achieved more than a 1 percentage point increase in proficiency in reading and math for blacks between 1998 and 1999. A slightly higher percentage achieved at least a 1 percentage point increase in reading and math for whites (76 percent).[6] However, only about half of these schools (55.8 percent) achieved more than a more than 1 percentage point increase in both subjects for *both* blacks and whites in that year. Ironically, if the increases for each were largely due to random fluctuations and were independent, we would have expected a very similar proportion (52.7 percent) to have achieved such growth for both groups (.693*.761=.527).

The odds are even longer for schools containing 3 racial or ethnic subgroups. Of the schools with more than 15 students in each of three racial groups-- blacks, whites and Hispanics-- only 26 percent achieved adequate yearly progress for *all three* groups. Again, this is only slightly more than we would have predicted if the changes for all groups were largely independent. Among these schools, 56 percent achieved adequate yearly progress for whites, 85 percent achieved adequate yearly progress for blacks and 44 percent achieved adequate yearly progress for Hispanics. If each group's scores were fluctuating independently, we would have expected only 21 percent to achieve adequate yearly progress for all three groups (.85*.44*.56=.21).

---

[6]The improvement in scores happened to be slightly larger between 1998 and 1999 for whites than for blacks. However, as reported in Table 4, the average annual increases over the period 1994 and 1999 were slightly larger for blacks at these schools.

Therefore, for purely statistical reasons, diverse schools are much less likely to achieve adequate yearly progress. Even if they are doing as well on average as all other schools in raising achievement for each of their racial subgroups, there is often a good chance that not all groups will see improvements in the same year.

Do the Subgroup Targets Actually Lead Schools to Focus on Performance of Minority Students?

Despite such unintended consequences, one might still wonder whether the use of subgroup targets actually spurs schools to focus on the performance of low-performing minority groups. In Texas, for example, in order for a school to achieve an "exemplary" rating, 90 percent of any racial and ethnic subgroup that represents more than 10 percent of the student body (and more than 30 students) must achieve proficiency. In other words, if a minority group represents less than 10 percent of the student body (for example, 9 percent), a school does not face a separate threshold for that group. However, if a minority group represents more than 10 percent of a school's students (for example, 11 percent), the school is held accountable for that group's performance separately (as long as there are also more than 30 students in the group).[7] In order to evaluate whether schools focus more on minority student performance as a result of such a rule, one could simply compare the change over time for minority students in schools where they represented more and less than 10 percent of the student body.

Figure 2 portrays the trend in the percent proficient for Latino students in schools where they represented 0 to 5 percent, 5 to 10 percent, 10 to 15 percent and more than 15 percent of the student body. (Latino students in schools where they had less than 30 Latino schoolmates were included with the 0 to 5 percent category.) Latino students in schools where they represented 5 to 10 percent of the student body had very similar levels of proficiency in 1994 as Latino students in schools where they made up 10 to 15 percent of the student body. (These two groups of schools are represented by the two middle lines in Figure 2.) Moreover, the trend over time was very similar. In other words, the improvement in performance for Latino students was unrelated to whether or not the school was being held accountable for Latino scores separately.

Figure 3 presents a similar figure for African American students. Again, the performance of African American students rose no more rapidly in schools where they were just above the threshold for separate consideration than for schools where they represented too small a share of the student body to be counted separately. As a result, there is very little evidence that creating an extra hurdle for schools led them to focus on minority student performance any more than in schools which faced no such extra hurdle.

---

[7]These rules apparently have a large impact on the proportion of schools achieving exemplary status. For example, elementary schools that were 5 to 10 percent Latino were three times as likely to achieve exemplary status than schools that were 10 to 15 percent Latino (32 percent versus 9 percent).

3. Conclusion and Recommendations

Whether designed by individual states or imposed by federal mandate, all school accountability systems face several fundamental challenges. First, they must determine how to measure a school's performance. Some, such as North Carolina, try to measure a school's value-added, by focusing upon the average improvement in student performance over the course of each grade. Under such a system, schools which enroll students who are underperforming in the early grades are put on a level playing field with schools whose students initially enroll better prepared. Others, such as California, base rewards and sanctions on the change in student performance from one calendar year to the next. Such an approach avoids the question of whether or not some schools face a tougher chore than other schools and, instead, tries to give all schools an incentive to improve. Still others, such as Texas, largely base their ratings on the average absolute level of performance of their students.

Second, a decision must be made regarding the time period upon which to base the assessment. We believe that most states place an inordinate amount of weight on the most recent year's worth of test score data. Several states are belatedly coming to recognize the costs of doing so. (See Kane and Staiger (2001 and forthcoming, 2002).)

A third challenge is presented by the large differences in test performance by racial and ethnic group. On one hand, the designers of accountability systems must be careful not to simply accept longstanding differences in performance by race and permanently lower expectations for minority youth; on the other hand, schools that serve disadvantaged minority youth must not be placed at such a disadvantage that they come to believe that success is out of reach.

Thus far, no consensus has emerged regarding any single best way to design school accountability systems. Until then, it would not be prudent to enshrine any particular approach into federal law. We would make the following three specific suggestions to the House and Senate conferees who will be working to resolve their differences in the coming weeks:

1. **Pool Performance over Multiple Years:** No serious consequences should be attached to one year of test score data because single years are so unreliable. Both the House and Senate bills would generate unnecessary paperwork, requiring schools to produce school improvement plans based upon single-year fluctuations in test scores. Any definition of adequate yearly progress should be based upon multiple years of performance data. Those schools that are not meeting expectations or making adequate yearly progress over 5 years should face serious consequences. However, the intermediate steps in both the House and Senate bills for schools that fail to make adequate progress for one year or two years will often be undeserved and, as such, may actually distract schools from the task of meeting their longer term objective.

At the end of 5 years, state governments should be required to certify to their citizens and to the federal government which schools have met expectations and which schools are making adequate progress. Schools identified by states to have failed to make adequate progress over that time period should face the serious consequences spelled out in the proposed legislation: reconstitution, public school choice, funding for supplemental education expenses.

2. **Maintain State Flexibility to Define Adequate Yearly Progress:** States should be free to define adequate progress in a manner that is consistent with the accountability systems they have been designing. For example, North Carolina should be allowed to define adequate progress in terms of the value-added composites they have been using to rate schools since 1997, rather than the percentage growth in proficiency written into the current federal legislation. Likewise, California should be allowed to use changes in its Academic Performance Index to rate school improvement. For example, 78 percent of the schools rated "exemplary" by the state of Texas in 1998 would have failed to make adequate yearly progress under the House rules. As long as state policymakers are required to report to their citizens the test performance of each school on an annual basis and as long as they are willing to certify to the same citizens which schools are making adequate progress and which are not, they should not be required to send mixed signals to schools, rating them on one measure for state purposes and rating them on another measure to satisfy the federal government.

   If, for any reason, federal policy makers fear that state government are unprepared to identify schools needing more serious intervention, the federal law could require that any such definition of adequate progress capture some minimum percentage of schools. For instance, the states scoring in the bottom quartile of the National Assessment of Educational Progress (or some other nationally-normed test) could be required to identify at least 20 percent of schools as having failed to make adequate progress, with states in the top three quartiles required to identify a smaller share of their schools as needing improvement.

   States should be given the flexibility to experiment with alternative ways to pool student performance data over multiple years. Some states may choose to use average improvement over several years. There are also more sophisticated ways to pool information over time which we propose in Kane and Staiger (2001). The federal law should allow states to experiment with different methods for averaging data.

3. **Do Not Penalize Racially Diverse Schools:** States should be required to report subgroup test performance, including by race and ethnicity, at the school level. However, sanctions should be imposed only when there is sufficient evidence that some racial groups are being left behind. The current legislation would do so in a haphazard manner. The more racial

subgroups any school contains it is much more likely to fail the current standards, for simple statistical reasons. Ironically, to the extent that disadvantaged minority students are more likely to attend racially diverse rather than racially homogeneous schools, such measures may end up hurting, rather than helping disadvantaged minority youth.

Over the next year, the U.S. Department of Education should be required to propose a method for identifying schools where there is sufficient evidence of divergent improvements in performance by race. States should then be prohibited from certifying any such schools as having achieved adequate progress at the end of five years.

There are real differences in performance at the school level. And schools that are not improving should be identified for intervention. However, one year's worth of test score data is insufficient to discern such differences in a meaningful way. Moreover, states are currently experimenting with a wide range of different types of accountability systems. They should be allowed to continue experimenting, until the Nation reaches a consensus regarding the ideal way to determine which schools are making adequate yearly progress and which schools are not. We understand the impulse to create a system which requires specific remedies sooner rather than later. However, impatience is an insufficient excuse for bad education policy. The current debate over the Elementary and Secondary Education Act could add momentum to ongoing state efforts to construct coherent accountability systems or it could generate a new set of distractions for schools and school districts. The suggestions outlined above are intended to ensure that ongoing state efforts at school reform stay on track.

References:

Kane, Thomas J. and Douglas O. Staiger "Improving School Accountability Measures" National Bureau of Economic Research Working Paper No. 8156, March 2001.

Kane, Thomas J. and Douglas O. Staiger "Volatility in School Test Scores: Implications for Test-Based Accountability Systems" forthcoming in Diane Ravitch (ed.) *Brookings Papers on Education Policy, 2002* (Washington, DC: Brookings Institution, forthcoming 2002).

Table 1.
Number of Years that North Carolina Elementary Schools Experienced Positive Improvments in Proficiency between 1994 and 1999

# of Years of Positive Change in Proficiency:	Percent of NC Elementary Schools:		
	In Math	In Reading	In Reading and In Math
0	.1	0	1.4
1	1.9	1.7	11.5
2	11.1	21.1	33.7
3	39.4	41.0	36.3
4	36.6	30.0	15.4
5	11.1	6.3	1.7
Average:	3.4	3.2	2.6

Note: Based upon authors' tabulations of the 1994-1999 NC end-of-grade test scores in grades 3 through 5 for 1023 schools in North Carolina that had students in all three grades in every year between 1994 and 1999. Students scoring at levels III and IV in reading or math were considered proficient.

Table 2.
Number of Years North Carolina Elementary Schools Achieved "Adequate Yearly Progress" between 1994 and 1999 Using the Definitions in the House and Senate Bills

# of Years Achieving Adequate Yearly Progress	House Bill	Senate Bill	
		Without Averaging	With 3-Year Rolling Average
0	48.4	25.6	26.8
1	37.4	35.9	25.4
2	12.2	27.1	21.7
3	1.9	9.5	16.0
4	.1	2.0	8.6
5	0	0	1.5
% Failing 1 or More Years (Must submit School Improv. Plan)	100	100	98
% Requ to Offer Public Sch Choice	100	88	80
% Facing Corrective Action	97	62	61
% Facing Restructuring	83	26	27

Note: Based upon authors' tabulations of the 1994-1999 NC end-of-grade test scores in grades 3 through 5. Students scoring at levels III and IV were considered proficient. The Senate rules require a 1 percentage point rise in the proportion of students proficient in reading and math in each subgroup. The House rules require a rise in the proportion of students proficient in each subgroup as well as at the school level to keep the school and each subgroup on track to achieve 100 percent proficiency in 12 years. Our data allowed us to identify up to 6 subgroups within each school: African American (non-Hispanic) students, Asians, Native Americans, Hispanics, white non-Hispanics and students with limited English proficiency. Only those subgroups consisting of 15 or more students were considered separately. Public school choice must be offered after 1 year of failure under the House bill and 2 consecutive years of failure under the Senate bill. Corrective actions are required after two consecutive years of failure under the House bill and after 3 consecutive years under the Senate bill. Corrective actions may involve: replacing relevant school staff, implementing a new curriculum and training teachers, increasing district oversight over school management, appointing experts to advise the school on its progress toward AYP, or extending the school year or day. Restructuring is required at the end of 3 years under the House bill and at the end of 5 years under the Senate Bill. Restructuring may involve conversion to a charter school, replacing the principal and most staff, contracting with a private entity to manage the school, or turning the operation of the school over to the state.

BEST COPY AVAILABLE

Table 3.
Number of Years Texas Elementary Schools
Achieved "Adequate Yearly Progress" between 1994 and 1999
Using the Definitions in the House and Senate Bills

# of Years Achieving Adequate Yearly Progress	House Bill	Senate Bill	
		Without Averaging	With 3-Year Rolling Average
0	35.0	27.5	22.0
1	44.3	40.5	24.6
2	17.5	23.0	24.3
3	3.0	7.6	17.6
4	.2	1.4	9.1
5	0	0	2.4
% Failing 1 or More Years (Must submit School Improv. Plan)	100	100	97
% Requ to Offer Public Sch Choice	100	90	78
% Facing Corrective Action	96	63	54
% Facing Restructuring	73	27	22

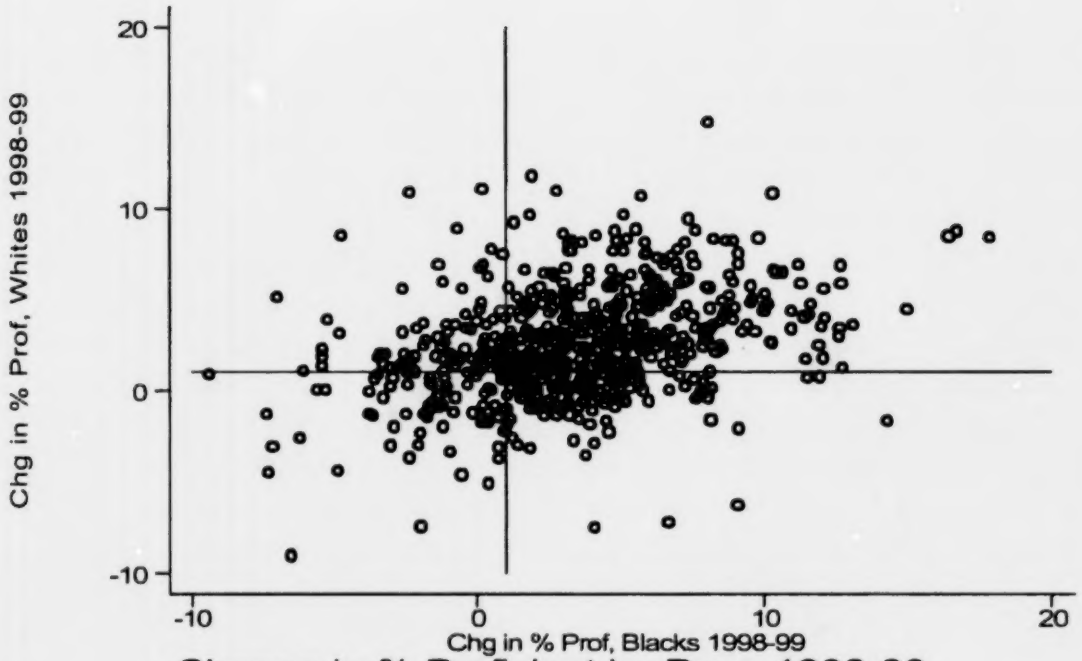
Note: Based upon authors' tabulations of the 1994-1999 TX test scores in grades 3 through 5. The Senate rules require a 1 percentage point rise in the proportion of students proficient in reading and math in each subgroup. The House rules require a rise in the proportion of students proficient in each subgroup as well as at the school level to keep the school and each subgroup on track to achieve 100 percent proficiency in 12 years. Our data allowed us to identify up to 4 subgroups within each school: African American (non-Hispanic) students, Hispanics, white non-Hispanics and students from "economically disadvantaged" backgrounds. Only those subgroups consisting of 15 or more students were considered separately. Public school choice must be offered after 1 year of failure under the House bill and 2 consecutive years of failure under the Senate bill. Corrective actions are required after two consecutive years of failure under the House bill and after 3 consecutive years under the Senate bill. Corrective actions may involve: replacing relevant school staff, implementing a new curriculum and training teachers, increasing district oversight over school management, appointing experts to advise the school on its progress toward AYP, or extending the school year or day. Restructuring is required at the end of 3 years under the House bill and at the end of 5 years under the Senate Bill. Restructuring may involve conversion to a charter school, replacing the principal and most staff, contracting with a private entity to manage the school, or turning the operation of the school over to the state.

Table 4.
Implications of the Subgroup Rules for Diverse Schools
(Using 3-Year Rolling Averages under Senate Rules, 1997-99)

	1 Subgroup	2 Subgroups	3 Subgroups
	Under the Senate Rules, using 3-Year Rolling Averages:		
Average Annual Proportion Achieving "Adequate Yearly Progress" 1997-99	.56	.34	.15
% Ever Failing in 3 Years	73	89	100
% Failing 3 Consecutive Years	15	38	64
	With Schoolwide Target Only and No Subgroup Rules		
Average Annual Proportion Achieving "Adequate Yearly Progress" 1997-99	.60	.54	.45
% Ever Failing in 3 Years	68	74	81
% Failing 3 Consecutive Years	13	20	26
	Racial/Ethnic Difference in Average Annual Growth in Math Proficiency		
White - African American	---	-.014	-.008
White - Latino	---	-.005	-.005
# of Schools:	334 (32.7%)	656 (64.2%)	31 (3.0%)

Note: Based upon authors' tabulations of the 1994-1999 NC end-of-grade test scores in grades 3 through 5. Students scoring at levels III and IV were considered proficient. The Senate rules require a 1 percentage point rise in the proportion of students proficient in reading and in math in each subgroup. Our data allowed us to identify up to 6 subgroups within each school: African American (non-Hispanic) students, Asians, Native Americans, Hispanics, white non-Hispanics and students with limited English proficiency. Only those subgroups consisting of 15 or more students in a school were considered separately.

Figure 1.



Change in % Proficient by Race 1998-99

	Achieved Adequate Progress for Whites:		
Achieved Adequate Progress for Blacks:	No	Yes	Row Total:
No	.104	.202	.307
Yes	.134	.558 (.693*.761=.527)	.693
Column Total:	.239	.761	1.00

Figure 2.

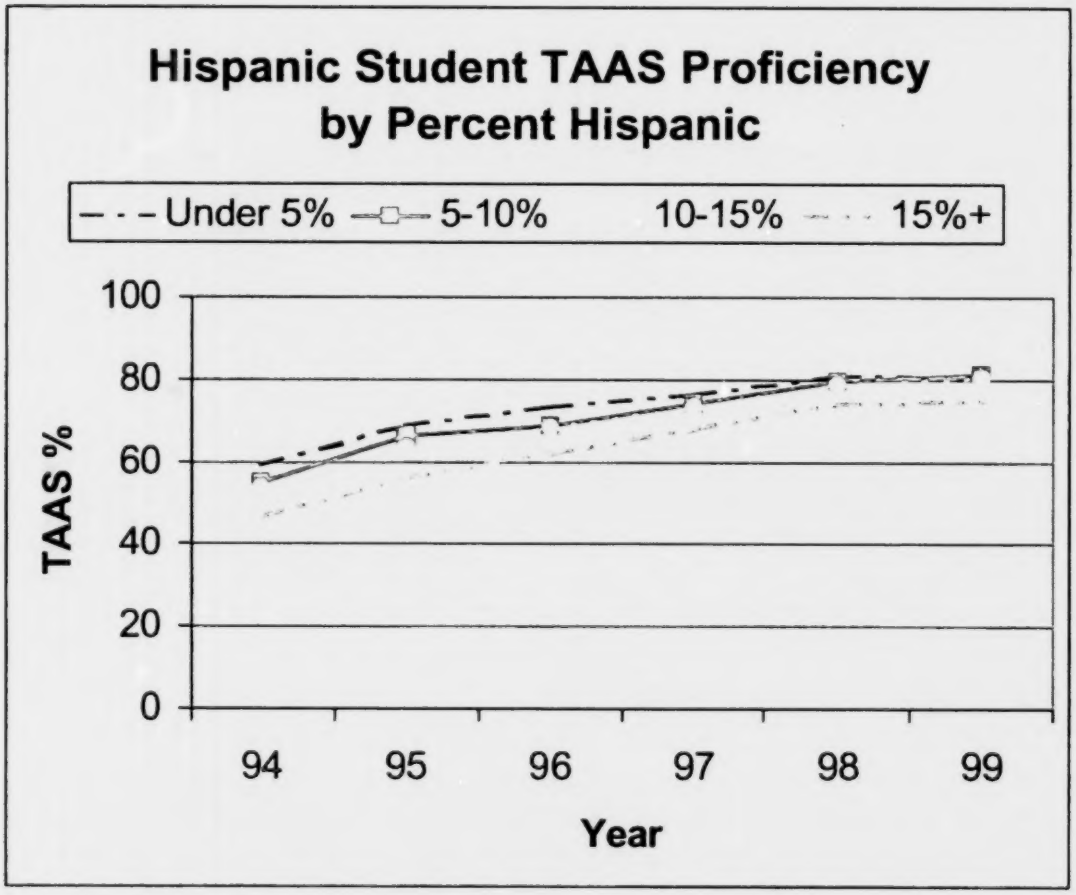
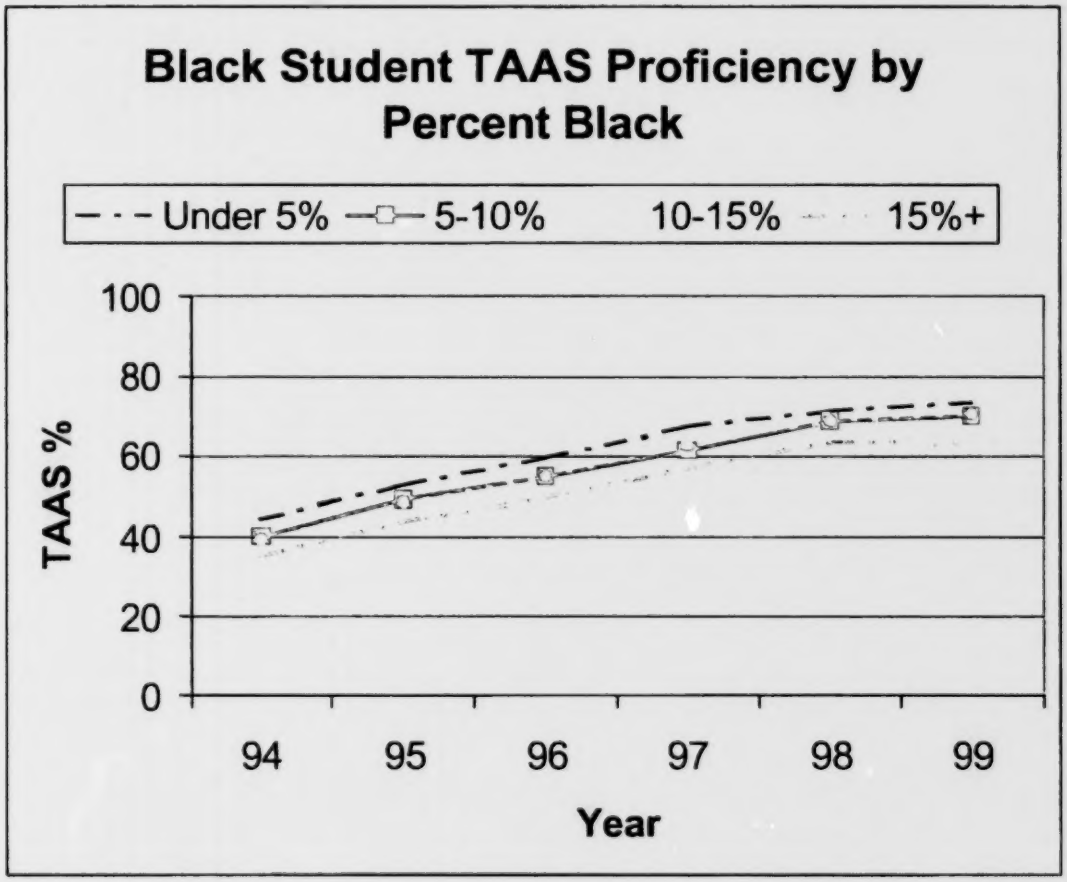

Hispanic Student TAAS Proficiency
by Percent Hispanic
Under 5%
5-10%
10-15%
15%+
100
80
60
40
20
0
TAAS %
94
95
96
97
98
99
Year

Figure 3.


Black Student TAAS Proficiency by Percent Black
Under 5%
5-10%
10-15%
15%+
TAAS %
100
80
60
40
20
0
94
95
96
97
98
99
Year

Appendix

The Senate bill would allow states to average the data for any given year with the data for the previous 2 years when calculating "adequate yearly progress." However, we were not able to calculate a 3 year average until 1996. If the law were implemented next year, some states will have historical data and would be able to start with a 3-year rolling average in the first year. In Table 2, we used a single year for 1994, averaged 1994 and 1995 to calculate the 1995 score and used the 3-year rolling average for the remaining years. In Appendix Table 1, we report the number of years schools achieved "adequate yearly progress" between 1997 and 1999-- three years for which we would have been able to calculate 3-year rolling averages for all schools.

Appendix Table 1.
Number of Years North Carolina Elementary Schools Achieved "Adequate Yearly Progress" between 1997 and 1999 Using the Definitions in the House and Senate Bills

# of Years Achieving Adequate Yearly Progress	House Bill	Senate Bill	
		Without Averaging	With 3-Year Rolling Average
0	62.8	40.6	31.2
1	31.5	41.5	31.1
2	5.5	15.7	22.2
3	0.3	2.2	15.5
% Failing 1 or More Years (Must submit School Improv. Plan)	100	98	84
% Requ to Offer Public Sch Choice	100	62	52
% Facing Corrective Action	80	41	31
% Facing Restructuring	63	(Requires 5 years of data)	(Requires 5 years of data)

Note: Based upon authors' tabulations of the 1994-1999 NC end-of-grade test scores in grades 3 through 5. Students scoring at levels III and IV were considered proficient. The Senate rules require a 1 percentage point rise in the proportion of students proficient in reading and math in each subgroup. The House rules require a rise in the proportion of students proficient in each subgroup as well as at the school level to keep the school and each subgroup on track to achieve 100 percent proficiency in 12 years. Our data allowed us to identify up to 6 subgroups within each school: African American (non-Hispanic) students, Asians, Native Americans, Hispanics, white non-Hispanics and students with limited English proficiency. Only those subgroups consisting of 15 or more students were considered separately. Public school choice must be offered after 1 year of failure under the House bill and 2 consecutive years of failure under the Senate bill. Corrective actions are required after two consecutive years of failure under the House bill and after 3 consecutive years under the Senate bill. Corrective actions may involve: replacing relevant school staff, implementing a new curriculum and training teachers, increasing district oversight over school management, appointing experts to advise the school on its progress toward AYP, or extending the school year or day. Restructuring is required at the end of 3 years under the House bill and at the end of 5 years under the Senate Bill. Restructuring may involve conversion to a charter school, replacing the principal and most staff, contracting with a private entity to manage the school, or turning the operation of the school over to the state.

BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

UD 035 592

I. DOCUMENT IDENTIFICATION:

Title:	"ASSESSING THE DEFINITION OF ADEQUATE YEARLY PROGRESS IN THE HOUSE AND SENATE EDUCATION BILLS"
Author(s):	THOMAS J. KANE, DOUGLAS STAIGER & JEFFREY GEPPERT
Corporate Source:	Publication Date: JULY, 2001

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to each document.

If permission is granted to reproduce and disseminate the identified documents, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

SAMPLE

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1 ☒

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

SAMPLE

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A ☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

SAMPLE

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B ☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate these documents as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature: Thomas J. Kane	Printed Name/Position/Title: PROFESSOR	
Organization/Address: UCLA	Telephone: 310-825-9413	FAX: 310-206-0337
	E-Mail Address: TOMKANE@UCLA.EDU	Date: